

AN APPLICATION OF SOURCE INEQUALITIES FOR CONVERGENCE RATES OF TIKHONOV REGULARIZATION WITH A NON-DIFFERENTIABLE OPERATOR

MARKUS GRASMAIR

ABSTRACT. In this paper we study Tikhonov regularization for the stable solution of an ill-posed non-linear operator equation. The operator we consider, which is related to an active contour model for image segmentation, is continuous, compact, but nowhere differentiable. Nevertheless we are able to derive convergence rates under different smoothness assumptions on the true solution by employing the method of variational or source inequalities. With this approach, we can prove up to linear convergence with respect to the norm.

KEYWORDS. Tikhonov regularization; convergence rates; source inequalities; non-linear ill-posed problems; active contours.

AMS SUBJECT CLASSIFICATIONS: 65J20; 65J22;

1. INTRODUCTION

The study of convergence rates within the context of solutions of ill-posed problems is concerned with the question of how good one can approximate the true solution of an equation given some noisy data of known noise level δ . Assume that $A: X \rightarrow Y$ is a bounded linear operator between two Hilbert spaces X and Y and we are trying to solve the linear operator equation $Ax = y^\delta$ with noisy right hand side y^δ satisfying $\|y - y^\delta\| \leq \delta$ for some “true” data $y \in Y$. Then the classical Tikhonov approach consists in the minimization of the functional

$$(1) \quad \|Ax - y^\delta\|_Y^2 + \alpha \|x\|_X^2$$

with some regularization parameter $\alpha > 0$. If δ and α tend to zero in such a way that $\delta^2/\alpha \rightarrow 0$, then the minimizers x_α^δ of (1) will converge to the “true” solution

$$x^\dagger = A^\dagger y$$

with A^\dagger being the Moore–Penrose inverse of A [11]. Moreover, it is possible to estimate the quality of the approximation, that is, the difference $\|x_\alpha^\delta - x^\dagger\|$, provided one has additional a-priori knowledge about smoothness properties of x^\dagger : If $x^\dagger = (A^*A)^\nu \omega$ for some $\omega \in X$ and $0 < \nu \leq 1$, then one can show for the parameter choice $\alpha \sim \delta^{\frac{2}{2\nu+1}}$ that we have [11, Corollary 3.1.4]

$$(2) \quad \|x_\alpha^\delta - x^\dagger\| = O(\delta^{\frac{2\nu}{2\nu+1}}) \quad \text{as } \delta \rightarrow 0.$$

In the case of a non-linear operator equation $F(x) = y$, the classical approach to the derivation of convergence rates uses a linearization of F at the true solution x^\dagger . Then, analogous convergence rates as in the linear case are derived under the assumption of a source condition of the form

$$(3) \quad (F'(x^\dagger)^* F'(x^\dagger))^\nu \omega = x^\dagger.$$

In the proofs of these convergence rates, however, it is necessary to introduce additional assumptions coupling the non-linear functional F and its linearization at x^\dagger .

An alternative approach to non-linear problems has been suggested in [16]. Instead of linearizing the problem and using source equalities, they proposed to use a single *source inequality* of the form

$$(4) \quad \|x - x^\dagger\|_X^2 \leq \|x\|_X^2 - \|x^\dagger\|_X^2 + \beta_2 \|F(x) - F(x^\dagger)\|_Y.$$

With this assumption, they were able to prove a convergence rate in the norm of order $O(\sqrt{\delta})$. This corresponds to the case $\nu = 1/2$ in (2); indeed, in the linear case, the source inequality (4) is equivalent to the source condition $x^\dagger = (A^*A)^{\frac{1}{2}}\omega$ (see [16]). Moreover, the approach has been generalized in [2, 7] to intermediate convergence rates, roughly corresponding to the source condition $x^\dagger = (A^*A)^\nu\omega$ with $0 \leq \nu \leq 1/2$.

Apart from the possibility of treating non-linear operators, source inequalities have the advantage that they do not require a Hilbert space structure of the set X ; indeed, the results in [16] were formulated in a Banach space setting ([7] does not even require linear spaces). One problem in Banach spaces is that fractional powers of a linear operator are not defined any more; a source condition of the form (3) cannot even be formulated. There is, however, a different approach to convergence rates of general order for non-linear problems in Banach spaces by means of *approximate source conditions* (see [12, 13, 14, 15]). Still this approach requires a prior linearization of the operator.

While theoretically the approach by source inequalities can be used for deriving convergence rates in general non-smooth situations, most cases where it is been applied were differentiable settings with a weak coupling between the operator and its linearization. In this paper we will show by means of a concrete example that the method of source inequalities can also be used for deriving convergence rates under meaningful conditions for a non-differentiable operator. We consider a situation that is closely related to an active contour model for image segmentation, where the operator, in its natural Hilbert space setting, is Hölder continuous but nowhere differentiable. Nevertheless it is possible to derive convergence rates up to order $O(\delta)$ under the assumption of sufficient smoothness of the true solution (in the sense that it lies in a certain Sobolev space).

2. SETTING

Denote by $H_+^1(S^1)$ the non-negative cone in the Sobolev space $H^1(S^1)$. That is, $\gamma \in H_+^1(S^1)$ if and only if $\gamma \in H^1(S^1)$ and $\gamma(t) \geq 0$ for every $t \in S^1$. If $\gamma \in H_+^1(S^1)$, we denote by

$$\text{hyp}(\gamma) := \{(t, x) \in S^1 \times \mathbb{R}_{\geq 0} : 0 \leq x \leq \gamma(t)\}$$

the *hypograph* of γ . That is, $\text{hyp}(\gamma)$ consists of all points lying between the t -axis and the graph of γ . We now define the mapping $F: H_+^1(S^1) \rightarrow L^2(S^1 \times \mathbb{R}_{\geq 0})$,

$$F(u) := \chi_{\text{hyp}(\gamma)}$$

with

$$\chi_{\text{hyp}(\gamma)}(t, x) := \begin{cases} 0 & \text{if } \gamma(t) < x, \\ 1 & \text{if } 0 \leq \gamma(t) \leq x, \end{cases}$$

which maps the curve γ to the characteristic function of its hypograph.

Note that every function $\gamma \in H_+^1(S^1)$ is bounded and continuous, and therefore $\text{hyp}(\gamma) \subset S^1 \times \mathbb{R}_{\geq 0}$ is a (non-empty) compact set. In particular this implies that the mapping F indeed takes values in the space $L^2(S^1 \times \mathbb{R}_{\geq 0})$. Moreover it is obvious that F is injective.

We now consider the problem of solving the equation $F(\gamma) = u$ for some given $u \in \text{Ran}(F)$ in a stable way. In addition, we assume that, instead of u , we are only given noisy data $u^\delta \in L^2(S^1 \times \mathbb{R}_{\geq 0})$ satisfying $\|u^\delta - u\|_{L^2} \leq \delta$, where $\delta > 0$ is some known noise level. Here we do not assume that the noisy data lie in the range of F or even that they are the characteristic function of some (Borel) set.

In order to find an approximate solution of the equation

$$(5) \quad F(\gamma) = u^\delta,$$

we apply classical (quadratic) Tikhonov regularization using the squared homogeneous first order Sobolev norm as regularization term. That is, we consider the Tikhonov functional

$$\mathcal{T}_\alpha(\gamma; u^\delta) := \|F(\gamma) - u^\delta\|_{L^2}^2 + \alpha \|\dot{\gamma}\|_{L^2}^2$$

and define, for $\alpha > 0$, an approximate solution of (5) by

$$\gamma_\alpha^\delta := \arg \min \{ \mathcal{T}_\alpha(\gamma; u^\delta) : \gamma \in H_+^1(S^1) \}.$$

Relation to Active Contours. One of the basic problems in image processing is segmentation: Given a possibly noisy image, interpreted as a function $u \in L^2(\mathbb{R}^2)$, the task in its simplest form is to partition this image into two regions, one representing objects in the image, the other representing the background. To that end, Chan and Vese [3] proposed a variational method based on the assumption that the objects (which are assumed to form a simply connected region) share a common grey value c_1 , and also the background has a uniform grey value $c_2 \neq c_1$. Without loss of generality we may assume that $c_2 = 1$ and $c_1 = 0$. Then the method proposed in [3] consists in minimizing, for some regularization parameters $\alpha_1, \alpha_2 > 0$, the functional

$$(6) \quad I(\gamma) := \alpha_1 \text{length}(\gamma) + \alpha_2 \text{area}(\gamma) + \int_{\text{int}(\gamma)} (1 - u(x))^2 dx + \int_{\mathbb{R}^2 \setminus \text{int}(\gamma)} u(x)^2 dx.$$

Here $\text{length}(\gamma)$ denotes the length of the curve γ , $\text{int}(\gamma) \subset \mathbb{R}^2$ denotes the set of all points lying inside the curve γ , and $\text{area}(\gamma)$ the size of the area inside the curve. This functional is minimized over the space $\text{Emb}(S^1; \mathbb{R}^2)$ of all closed curves continuously embedded in \mathbb{R}^2 .

Defining the function $F: \text{Emb}(S^1; \mathbb{R}^2) \rightarrow L^2(\mathbb{R}^2)$,

$$F(\gamma) := \chi_{\text{int}(\gamma)},$$

one can write the functional equivalently as

$$I(\gamma) = \alpha_1 \text{length}(\gamma) + \alpha_2 \text{area}(\gamma) + \|F(\gamma) - u\|_{L^2}^2.$$

Moreover, for γ sufficiently smooth, one can write

$$\text{length}(\gamma) = \int_{S^1} |\dot{\gamma}(t)| dt, \quad \text{area}(\gamma) = \frac{1}{2} \int_{S^1} \langle \gamma(t), \dot{\gamma}(t) \rangle dt.$$

Thus the Chan–Vese model (6) can be brought in a quite similar form as the problem we consider in this paper, which, conversely, can be interpreted as a simplified active contour model for images on a half-cylinder. A major reason for our

simplification is the fact that it allows us to work in a Hilbert space setting. In contrast, the “natural” setting for the Chan–Vese model would rather be the Banach manifold of Lipschitz embeddings of S^1 in \mathbb{R}^2 modulo reparameterizations of S^1 (note that $I(\gamma)$ is invariant with respect to reparameterizations of γ ; in order to obtain any uniqueness result it is therefore necessary to factor out reparameterizations).

3. WELL-POSEDNESS OF THE REGULARIZATION METHOD

We now prove that quadratic Tikhonov regularization yields a well-posed regularization method for the functional F . To that end we mainly have to investigate the continuity properties of F .

Lemma 3.1. *The mapping F is Hölder continuous of degree $1/2$ as a mapping from $L_+^2(S^1)$ to $L^2(S^1 \times \mathbb{R}_{\geq 0})$. In particular, it is weakly continuous and compact as a mapping from $H_+^1(S^1)$ to $L^2(S^1 \times \mathbb{R}_{\geq 0})$.*

Proof. Let $\gamma_1, \gamma_2 \in L_+^2(S^1)$. Then

$$\begin{aligned} \|F(\gamma_1) - F(\gamma_2)\|_{L^2}^2 &= \int_{S^1} \int_{\mathbb{R}_{\geq 0}} |\chi_{\text{hyp}(\gamma_1)}(t, x) - \chi_{\text{hyp}(\gamma_2)}(t, x)|^2 dx dt \\ &= \int_{S^1} |\gamma_1(t) - \gamma_2(t)| dt = \|\gamma_1 - \gamma_2\|_{L^1} \leq \sqrt{2\pi} \|\gamma_1 - \gamma_2\|_{L^2}, \end{aligned}$$

which proves the Hölder continuity of F as a mapping from $L_+^2(S^1)$. The weak continuity and compactness of F as a mapping from $H_+^1(S^1)$ now follow from the fact that $H^1(S^1)$ is compactly embedded in $L^2(S^1)$. \square

Theorem 3.2. *The following hold:*

- (1) *For every $u \in L^2(S^1 \times \mathbb{R}_{\geq 0})$ and $\alpha > 0$ the functional $\mathcal{T}_\alpha(\cdot; u)$ attains its minimum in $H_+^1(S^1)$.*
- (2) *Assume that $u \in L^2(S^1 \times \mathbb{R}_{\geq 0})$ and $\alpha > 0$. Let $\{u^{(k)}\}_{k \in \mathbb{N}} \subset L^2(S^1 \times \mathbb{R}_{\geq 0})$ be any sequence converging to u , and let*

$$\gamma^{(k)} \in \arg \min \{ \mathcal{T}_\alpha(\gamma; u^{(k)}) : \gamma \in H_+^1(S^1) \}.$$

Then there exists a sub-sequence $\{\gamma^{(k_j)}\}_{j \in \mathbb{N}}$ converging strongly in $H^1(S^1)$ to some

$$\gamma \in \arg \min \{ \mathcal{T}_\alpha(\gamma; u) : \gamma \in H_+^1(S^1) \}.$$

- (3) *Assume that $u = F(\gamma^\dagger)$ for some $\gamma^\dagger \in H_+^1(S^1)$. Let $\{\delta_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_{>0}$ and $\{\alpha_k\}_{k \in \mathbb{N}} \subset \mathbb{R}_{>0}$ be sequences satisfying $\delta_k \rightarrow 0$, $\alpha_k \rightarrow 0$, and $\delta_k^2/\alpha_k \rightarrow 0$, and let $u^{\delta_k} \in L^2(S^1 \times \mathbb{R}_{\geq 0})$, $k \in \mathbb{N}$, satisfy $\|u_k - u\|_{L^2} \leq \delta_k$. Let*

$$\gamma^{(k)} := \gamma_{\alpha_k}^{\delta_k} \in \arg \min \{ \mathcal{T}_{\alpha_k}(\gamma; u^{\delta_k}) : \gamma \in H_+^1(S^1) \}.$$

Then the sequence $\{\gamma^{(k)}\}_{k \in \mathbb{N}}$ converges strongly in $H^1(S^1)$ to γ^\dagger .

Proof. This follows from standard results in the theory of (non-linear) Tikhonov regularization, which can be found, for instance, in [4, 19]. The only (ever so slight) complication is the fact that the regularization term uses the homogeneous Sobolev norm, which is not coercive on $H^1(S^1)$. The coercivity of the functional \mathcal{T}_α , however, follows immediately from the chain of inequalities

$$\|F(\gamma) - u\|_{L^2}^2 \leq 2\|F(\gamma)\|_{L^2}^2 + 2\|u\|_{L^2}^2 = 2\|\gamma\|_{L^1} + 2\|u\|_{L^2}^2 \leq 2\|\gamma\|_{L^2} + 2\|u\|_{L^2}^2.$$

□

Remark 1. Because F is injective, it follows that the equation $F(\gamma) = u$ can have at most one solution. Still it is possible that the optimisation problem $\mathcal{T}_\alpha(\gamma; u^\delta) \rightarrow \min$ can have multiple solutions, even for $\delta > 0$ arbitrarily small.

Consider for instance the situation where $\gamma^\dagger(t) = 1$ and $F(\gamma^\dagger) = \chi_{S^1 \times [0,1]}$. Define for $\delta > 0$ with $\delta^2 < \pi$ the function

$$u^\delta(t, x) := \begin{cases} 1 & \text{if } 0 \leq x < 1 - \delta^2/\pi, \\ \frac{1}{2} & \text{if } |x - 1| \leq \delta^2/\pi, \\ 0 & \text{if } x > 1 + \delta^2/\pi. \end{cases}$$

Then $\|u^\delta - u\|_{L^2} = \delta$.

Now let $\gamma \in H_+^1(S^1)$ and define

$$J^+(\delta) := \{t \in S^1 : \gamma(t) > 1 + \delta^2/\pi\}, \quad J^-(\delta) := \{t \in S^1 : \gamma(t) < 1 - \delta^2/\pi\}.$$

Then it is easy to see that

$$\|F(\gamma) - u^\delta\|_{L^2}^2 = \delta^2 + \int_{J^+(\delta)} \gamma(t) - 1 - \delta^2/\pi \, dt + \int_{J^-(\delta)} 1 - \gamma(t) - \delta^2/\pi \, dt.$$

Therefore $\|F(\gamma) - u^\delta\|_{L^2}^2$ is minimal (with value δ^2) if and only if $|\gamma(t) - 1| \leq \delta^2/\pi$ for every t . Thus it follows that

$$\mathcal{T}_\alpha(\gamma; u^\delta) \geq \delta^2$$

for every $\gamma \in H_+^1(S^1)$, and equality holds if and only if γ is a constant function of the form $\gamma(t) = 1 + c$ for some $|c| \leq \delta^2/\pi$. In other words, *every* such function is a minimizer of the Tikhonov functional with data u^δ and *any* regularization parameter $\alpha > 0$.

In particular, this example shows that the stability result in Theorem 3.2 (item (2)) really requires the formulation in terms of sub-sequences; it can happen that the sequence $\{\gamma^{(k)}\}_{k \in \mathbb{N}}$ itself does not converge. Note, however, that it is also possible to formulate the stability result in terms of set convergence. Such an approach has for instance been used in [10].

4. CONVERGENCE RATES

In this main section of the paper we will discuss the derivation of convergence rates, that is, quantitative estimates for the difference between the regularized solution γ_α^δ and the true solution γ^\dagger in dependence of α and δ .

Throughout this whole section we assume that the equation $F(\gamma) = u$ has a (necessarily unique) solution $\gamma^\dagger \in H_+^1(S^1)$. Moreover we denote by γ_α^δ *any* minimizer of the regularization functional $\mathcal{T}_\alpha(\cdot; u^\delta)$ for *any* noisy data $u^\delta \in L^2(S^1 \times \mathbb{R}_{\geq 0})$ satisfying $\|u^\delta - u\|_{L^2} \leq \delta$.

Classically, convergence rates for quadratic Tikhonov regularization of non-linear operators on Hilbert spaces have been derived for sufficiently smooth operators under the assumption that the true solution γ^\dagger is contained in the range of the adjoint of the derivative of the operator F at γ^\dagger . In addition it is necessary to compensate for the non-linearity of the operator F by assuming additional regularity properties of its derivative F' . One example is the following theorem taken from [19]. Similar results can also be found in [4, 5].

Theorem 4.1. *Assume that F is Gâteaux differentiable in a neighborhood of γ^\dagger and that there exist $\omega \in L^2(S^1 \times \mathbb{R}_{\geq 0})$ and $c > 0$ with $c\|\omega\|_{L^2} < 1$ such that*

$$\gamma^\dagger = F'(\gamma^\dagger)^*\omega$$

and

$$\|F(\gamma) - F(\gamma^\dagger) - F'(\gamma^\dagger)(\gamma - \gamma^\dagger)\|_{L^2} \leq \frac{c}{2}\|\gamma - \gamma^\dagger\|_{H^1}^2,$$

for every $\gamma \in H_+^1(S^1)$ sufficiently close to γ^\dagger . Then we have for a parameter choice $\alpha \sim \delta$ the rate

$$\|\gamma_\alpha^\delta - \gamma^\dagger\|_{H^1} = O(\sqrt{\delta}).$$

The crucial assumption in this theorem is the *source condition* $\gamma^\dagger \in F'(\gamma^\dagger)^*\omega$, which links the non-linear problem to the much better understood linear theory. The additional assumptions mainly guarantee that the gap between the non-linear and the linearized problem is sufficiently small. Still, the whole approach for the derivation of convergence rates relies on a linearization of the operator F , which in our situation is not easily possible, as the next result shows.

Lemma 4.2. *The mapping $F: H_+^1(S^1) \rightarrow L^2(S^1 \times \mathbb{R}_{\geq 0})$ is nowhere differentiable in the interior of its domain.*

Proof. First note that the interior of $H_+^1(S^1)$ consists of all functions $\gamma \in H^1(S^1)$ with $\gamma(t) > 0$ for every $t \in S^1$.

Let therefore $\gamma \in H_+^1(S^1)$ satisfy $\gamma(t) > 0$ everywhere, and let $\sigma \in H^1(S^1)$. Then, as shown in the proof of Lemma 3.1, we have

$$\|F(\gamma + s\sigma) - F(\gamma)\|_{L^2}^2 = \int_{S^1} |s\sigma(t)| dt.$$

Now assume that $\sigma \neq 0$. Then there exist a non-empty open subset $J \subset S^1$ and $\varepsilon > 0$ such that $|\sigma(t)| > \varepsilon$ for every $t \in J$. Consequently

$$\|F(\gamma + s\sigma) - F(\gamma)\|_{L^2}^2 = |s| \int_{S^1} |\sigma(t)| dt \geq \varepsilon |s| |J|.$$

Thus

$$\lim_{s \rightarrow 0^+} \frac{1}{s} \|F(\gamma + s\sigma) - F(\gamma)\|_{L^2} \geq \lim_{s \rightarrow 0^+} \sqrt{\frac{\varepsilon |J|}{s}} = +\infty,$$

which proves that F does not even have a one-sided directional derivative in any non-trivial direction. \square

For the next result—the main theorem of this paper—recall the definition of the fractional order Sobolev spaces $W^{s,q}$, $s \in \mathbb{R}$, $1 \leq q \leq \infty$, and the homogeneous Sobolev spaces $W_0^{s,q}$, which can for instance be found in [1].

Theorem 4.3. *Assume that $\gamma^\dagger \in W^{s,q}(S^1)$ with $1 < s \leq 2$ and $2 \leq q < +\infty$. Then there exist constants $c_1, c_2, c_3 > 0$ such that*

$$\begin{aligned} c_1 \|\gamma - \gamma^\dagger\|_{H_0^1}^2 \\ \leq \|\gamma\|_{H_0^1}^2 - \|\gamma^\dagger\|_{H_0^1}^2 + c_2 \|F(\gamma) - F(\gamma^\dagger)\|_{L^2}^{4-\frac{4}{s}} + c_3 \|F(\gamma) - F(\gamma^\dagger)\|_{L^2}^{2-\frac{2q}{qs-s+1}} \end{aligned}$$

for every $\gamma \in H_+^1(S^1)$.

If $\gamma^\dagger \in W^{s,\infty}(S^1)$ with $1 < s \leq 2$, then

$$(7) \quad c_1 \|\gamma - \gamma^\dagger\|_{H_0^1}^2 \leq \|\gamma\|_{H_0^1}^2 - \|\gamma^\dagger\|_{H_0^1}^2 + c_2 \|F(\gamma) - F(\gamma^\dagger)\|_{L^2}^{4-\frac{4}{s}}.$$

Proof. We only show the assertion for the more complicated case $q < \infty$.

First note that

$$\|\dot{\gamma} - \dot{\gamma}^\dagger\|^2 = \|\dot{\gamma}\|^2 - \|\dot{\gamma}^\dagger\|^2 - 2\langle \dot{\gamma}^\dagger, \dot{\gamma} - \dot{\gamma}^\dagger \rangle.$$

In order to estimate the last term on the right hand side of this equation, we will use the abbreviation

$$\sigma := \gamma - \gamma^\dagger.$$

Because $\gamma^\dagger \in W_+^{s,q}(S^1)$, it follows that

$$|\langle \dot{\gamma}^\dagger, \dot{\sigma} \rangle| \leq \|\gamma^\dagger\|_{W_0^{s,q}} \|\sigma\|_{W_0^{2-s,q*}}$$

Moreover the interpolation inequality for Sobolev functions implies that there exists some $C_1 > 0$ only depending on s and q_* such that

$$\|\sigma\|_{W_0^{2-s,q*}} \leq C_1 \|\sigma\|_{L^{q*}}^{s-1} \|\sigma\|_{W_0^{1,q*}}^{2-s} \leq C_1 \|\sigma\|_{L^{q*}}^{s-1} \|\sigma\|_{H_0^1}^{2-s}.$$

The last inequality holds, because $q \geq 2$ and therefore $q_* \leq 2$. Now note that

$$\begin{aligned} \|\sigma\|_{L^{q*}} &\leq \|\sigma\|_{L^1}^{\frac{1}{q_*}} \|\sigma\|_{L^\infty}^{\frac{1}{q}} \\ &\leq \|\sigma\|_{L^1}^{\frac{1}{q_*}} (\|\sigma\|_{L^1} + \|\sigma\|_{W_0^{1,1}})^{\frac{1}{q}} \\ &\leq \|\sigma\|_{L^1} + \|\sigma\|_{L^1}^{\frac{1}{q_*}} \|\sigma\|_{H_0^1}^{\frac{1}{q}}. \end{aligned}$$

Setting $C_2 := 2C_1 \|\gamma^\dagger\|_{W_0^{s,q}}$, we thus obtain the estimate

$$2|\langle \dot{\gamma}^\dagger, \dot{\sigma} \rangle| \leq C_2 \|\sigma\|_{L^1}^{s-1} \|\sigma\|_{H_0^1}^{2-s} + C_2 \|\sigma\|_{L^1}^{\frac{s-1}{q_*}} \|\sigma\|_{H_0^1}^{2-s+\frac{s-1}{q}}.$$

Now we use Young's inequality $ab \leq \frac{1}{p}a^p + \frac{1}{p_*}b^{p_*}$, which implies with $p = 2/s$ and $p_* = 2/(2-s)$ that

$$C_2 \|\sigma\|_{L^1}^{s-1} \|\sigma\|_{H_0^1}^{2-s} \leq \frac{s}{2} C_2^{\frac{2}{s}} \|\sigma\|_{L^1}^{\frac{2s-2}{s}} + \left(1 - \frac{s}{2}\right) \|\sigma\|_{H_0^1}^2$$

and with $p = \frac{2q}{qs-s+1}$ and $p_* = \frac{2q}{2q-qs+s-1}$

$$\begin{aligned} C_2 \|\sigma\|_{L^1}^{\frac{s-1}{q_*}} \|\sigma\|_{H_0^1}^{2-s+\frac{s-1}{q}} &\leq \frac{qs-s+1}{2q} C_2^{\frac{2q}{qs-s+1}} \|\sigma\|_{L^1}^{1-\frac{q}{qs-s+1}} \\ &\quad + \left(1 - \frac{qs-s+1}{2q}\right) \|\sigma\|_{H_0^1}^2. \end{aligned}$$

Thus we obtain with the constants

$$\begin{aligned} c_1 &= \frac{(2q-1)(s-1)}{2q}, \\ c_2 &= \frac{s}{2} C_2^{\frac{2}{s}}, \\ c_3 &= \frac{qs-s+1}{2q} C_2^{\frac{2q}{qs-s+1}}, \end{aligned}$$

the estimate

$$c_1 \|\gamma - \gamma^\dagger\|_{H_0^1}^2 \leq \|\gamma\|_{H_0^1}^2 - \|\gamma^\dagger\|_{H_0^1}^2 + c_2 \|\gamma - \gamma^\dagger\|_{L^1}^{2-\frac{2}{s}} + c_3 \|\gamma - \gamma^\dagger\|_{L^1}^{1-\frac{q}{qs-s+1}}.$$

Now the assertion follows from the fact that $\|F(\gamma) - F(\gamma^\dagger)\|_{L^2}^2 = \|\gamma - \gamma^\dagger\|_{L^1}^2$. \square

Corollary 4.4. *Let the conditions of Theorem 4.3 be satisfied and let*

$$\Phi(t) := c_2 t^{2\frac{s-1}{s}} + c_3 t^{\frac{s-1}{q_*}}$$

in case $q < \infty$ and

$$\Phi(t) := c_2 t^{2\frac{s-1}{s}}$$

in case $q = \infty$.

- (1) *If $q < \infty$ and $s < 2$ let Ψ be the convex conjugate of the strictly convex mapping $t \mapsto \Phi^{-1}(2t)$. Then*

$$c_1 \|\gamma_\alpha^\delta - \gamma^\dagger\|_{H_0^1}^2 \leq \frac{\delta^2}{\alpha} + \Phi(2\delta^2) + \frac{\Psi(\alpha)}{\alpha}.$$

- (2) *If $q = \infty$ and $s = 2$ we have*

$$c_1 \|\gamma_\alpha^\delta - \gamma^\dagger\|_{H_0^1}^2 \leq \frac{\delta^2}{\alpha} + 2c_2 \delta^2$$

whenever $\alpha \leq 1/(2c_2)$.

Proof. This is a direct consequence of Theorem 4.3 and [7, Theorem 3.1]. \square

Corollary 4.5. *Let the conditions of Theorem 4.3 be satisfied.*

- (1) *If $q < \infty$ or $s < 2$ we have for the parameter choice $\alpha \sim \delta^{2\frac{q_*-s+1}{q_*}}$ the convergence rate*

$$\|\gamma_\alpha^\delta - \gamma^\dagger\|_{H_0^1} = O(\delta^{\frac{s-1}{q_*}}).$$

- (2) *If $q = \infty$ and $s = 2$ we have for a constant parameter choice $\alpha \leq 1/(2c_2)$ the convergence rate*

$$\|\gamma_\alpha^\delta - \gamma^\dagger\|_{H_0^1} = O(\delta).$$

Proof. See [7, Corollary 3.1]. \square

Remark 2. Note that the claim of Theorem 4.3 in the case $s = 2$ and $q = \infty$ can be shown much more easily with elementary computations. Indeed, in this case one has

$$\begin{aligned} \|\gamma - \gamma^\dagger\|_{H_0^1}^2 &= \|\gamma\|_{H_0^1}^2 - \|\gamma^\dagger\|_{H_0^1}^2 - 2 \int_{S^1} \dot{\gamma}^\dagger(t) (\dot{\gamma}(t) - \dot{\gamma}^\dagger(t)) dt \\ &= \|\gamma\|_{H_0^1}^2 - \|\gamma^\dagger\|_{H_0^1}^2 + 2 \int_{S^1} \ddot{\gamma}^\dagger(t) (\gamma(t) - \gamma^\dagger(t)) dt \\ &\leq \|\gamma\|_{H_0^1}^2 - \|\gamma^\dagger\|_{H_0^1}^2 + 2 \|\ddot{\gamma}^\dagger\|_{L^\infty} \|\gamma - \gamma^\dagger\|_{L^1} \\ &= \|\gamma\|_{H_0^1}^2 - \|\gamma^\dagger\|_{H_0^1}^2 + 2 \|\ddot{\gamma}^\dagger\|_{L^\infty} \|F(\gamma) - F(\gamma^\dagger)\|_{L^2}^2. \end{aligned}$$

Thus (7) holds with $c_1 = 1$ and $c_2 = 2 \|\ddot{\gamma}^\dagger\|_{L^\infty}$.

5. FURTHER ASPECTS

5.1. Differentiability. In Lemma 4.2 we have shown that F is nowhere differentiable when regarded as a mapping from $H_+^1(S^1)$ to $L^2(S^1 \times \mathbb{R}_{\geq 0})$. We now discuss possible different settings, where F has better regularity properties.

Because the range of F consists only of characteristic functions with compact support, it follows that F can be regarded as a mapping into any space $L^p(S^1 \times \mathbb{R}_{\geq 0})$

with $1 \leq p \leq \infty$. Moreover the same computation as in the proof of Lemma 3.1 shows that for every $1 \leq p < \infty$ we have

$$\|F(\gamma_1) - F(\gamma_2)\|_{L^p}^p = \|\gamma_1 - \gamma_2\|_{L^1} \leq \sqrt{2\pi} \|\gamma_1 - \gamma_2\|_{L^2}.$$

Thus, seen as a mapping $F: L^2(S^1) \rightarrow L^p(S^1; \mathbb{R}_{\geq 0})$, the mapping F is Hölder continuous of degree $1/p$; in case $p = 1$, this shows that F is Lipschitz continuous. Note that for $p = \infty$, the mapping F is discontinuous everywhere. Concerning the differentiability of F , however, this change of the target space does not really matter. The same argumentation as in Lemma 4.2 shows that F is nowhere differentiable as a mapping into $L^p(S^1 \times \mathbb{R}_{\geq 0})$ with $1 < p < \infty$.

For $p = 1$ the situation is slightly different: If $\gamma \in H_+^1(I)$ satisfies $\gamma(t) > 0$ for every t and $\sigma \in H^1(I)$ is fixed, then the family of functions $\{\frac{1}{s}(F(\gamma + s\sigma) - F(\gamma))\}_{s>0}$ forms a bounded subset of $L^1(S^1 \times \mathbb{R}_{\geq 0})$. Still, the limit $\lim_{s \rightarrow 0^+} \frac{1}{s}(F(\gamma + s\sigma) - F(\gamma))$ does not exist in $L^1(S^1 \times \mathbb{R}_{\geq 0})$ and thus, again, F is nowhere directionally differentiable.

Finally, one can regard F as a mapping from $H_+^1(S^1)$ to the space $\mathcal{M}(S^1 \times \mathbb{R}_{\geq 0})$ of finite Radon measures on $S^1 \times \mathbb{R}_{\geq 0}$. Because $L^1(S^1 \times \mathbb{R}_{\geq 0})$ is isometrically embedded into $\mathcal{M}(S^1 \times \mathbb{R}_{\geq 0})$, this does not change the Lipschitz continuity of F . However, the difference quotients $\frac{1}{s}(F(\gamma + s\sigma) - F(\gamma))$ of F now have a limit in $\mathcal{M}(S^1 \times \mathbb{R}_{\geq 0})$ with respect to the weak* topology: One has

$$\frac{1}{s}(F(\gamma + s\sigma) - F(\gamma)) \rightharpoonup^* \gamma^\#(\sigma \mathcal{L}^1),$$

where $\gamma^\#(s\mathcal{L}^1)$ denotes the push forward of the one-dimensional Lebesgue measure \mathcal{L}^1 weighted by the functions σ via the mapping γ . That is,

$$\gamma^\#(\sigma \mathcal{L}^1)(h) = \int_{S^1} \sigma(t) h(t, \gamma(t)) dt$$

for every $h \in C_b(S^1 \times \mathbb{R}_{\geq 0})$.

The considerations above show that the setting of the problem can be modified in such a way that the operator F becomes differentiable. Thus it might still be possible to derive convergence rates in a more classical way via a linearization of F . There are, however, several difficulties: First, the extension to $\mathcal{M}(S^1 \times \mathbb{R}_{\geq 0})$ leaves the Hilbert space setting in favor of a more complicate Banach space setting (with non-reflexive spaces). In such a setting, non-standard convergence rates have been derived via linearization ideas in [14] by the technique of approximate source conditions. Second, even the Banach space setting might not be the correct one, as the operator F is only differentiable with respect to the weak* topology and not the norm topology. Thus it might even be necessary to derive results in general locally convex spaces. Finally note that after changing the topological structure, the similarity term is not the squared norm on the target space any more. Thus it might even be necessary to turn to results for convergence rates with more general similarity terms (see for instance [6, 7, 18]).

5.2. Differentiability of the Regularization Term. Interestingly, the regularization functional itself has potentially much better smoothness properties than F . In order to see this, define for $u \in L^2(S^1 \times \mathbb{R}_{\geq 0})$ the functional $\mathcal{S}: H_+^1(S^1) \rightarrow \mathbb{R}_{\geq 0}$,

$$\mathcal{S}_u(\gamma) := \|F(\gamma) - u\|_{L^2}^2.$$

Then it is easy to see that \mathcal{S}_u is Fréchet differentiable whenever $u \in L^2(S^1 \times \mathbb{R}_{\geq 0}) \cap C(S^1 \times \mathbb{R}_{\geq 0})$ with Fréchet derivative

$$\mathcal{S}'_u(\gamma) = \left[\sigma \mapsto \int_{S^1} \sigma(t) (1 - 2u(t, \gamma(t))) dt \right].$$

In addition, for $u = F(\gamma)$ the mapping \mathcal{S} , while not differentiable, possesses one-sided directional derivatives of the form

$$\mathcal{S}'_{F(\gamma)}(\gamma; \sigma) = 2 \int_{S^1} |\sigma(t)| dt = 2 \|\sigma\|_{L^1}.$$

The specific form of this one-sided derivative might possibly yield an explanation of the linear convergence rate that has been shown in Corollary 4.5 for the case where $\gamma^\dagger \in W^{2,\infty}(S^1)$, as such a rate is not possible for smooth operators using the squared Hilbert space norm as a regularization term. For quadratic regularization of Fréchet differentiable operators, it has been shown that the best possible convergence rate in non-trivial cases is of order $O(\delta^{2/3})$ (see [17]). Linear convergence rates, however, have been derived recently for Tikhonov regularization with non-smooth regularization terms. The first result in this direction was [9], where linear rates have been derived for regularization on the sequence space ℓ^2 with the ℓ^1 -norm as regularization term under the assumptions of sparsity of the true solution and a certain source condition. This result has also been extended in [8] to more general positively homogeneous regularization terms. The basis of all these results is the non-smoothness of the regularization term; the rates derived in this paper indicate that a sufficient non-smoothness of the operator F to be inverted can have similar effects.

6. CONCLUSION

We have shown in this paper that the approach of source inequalities for the derivation of convergence rates for Tikhonov regularization can be applied to non-linear problems where approaches based on linearization are bound to fail. Our paradigm was a functional related to the Chan–Vese active contour model for image segmentation, which was shown to be continuous but nowhere differentiable in its domain. Still, the approach by source inequalities allowed us to derive convergence rates under reasonable (and easily interpretable) smoothness assumptions on the true solution. Surprisingly, the convergence rates do not obey the classical bound of $O(\delta^{2/3})$, which is known to be best possible rate for quadratic regularization of differentiable functionals. Instead, for a sufficiently smooth true solution, we were able to obtain a rate of order $O(\delta)$. One possible explanation for this exceedingly good behaviour is a connection to sparse (or ℓ^1) regularization: For noise free data the similarity term in our problem has precisely the same behaviour as the ℓ^1 -regularization term.

REFERENCES

- [1] J. Bergh and J. Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.
- [2] R. I. Boţ and B. Hofmann. An extension of the variational inequality approach for obtaining convergence rates in regularization of nonlinear ill-posed problems. *J. Integral Equations Appl.*, 22(3):369–392, 2010.
- [3] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. Image Process.*, 10(2):266–277, 2001.

- [4] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
- [5] H. W. Engl, K. Kunisch, and A. Neubauer. Convergence rates for Tikhonov regularisation of nonlinear ill-posed problems. *Inverse Probl.*, 5(3):523–540, 1989.
- [6] J. Flemming. Theory and examples of variational regularization with non-metric fitting functionals. *J. Inverse Ill-Posed Probl.*, 18(6):677–699, 2010.
- [7] M. Grasmair. Generalized Bregman distances and convergence rates for non-convex regularization methods. *Inverse Probl.*, 26(11):115014, 2010.
- [8] M. Grasmair. Linear convergence rates for Tikhonov regularization with positively homogeneous functionals. *Inverse Probl.*, 27(7):075014, 2011.
- [9] M. Grasmair, M. Haltmeier, and O. Scherzer. Sparse regularization with l^q penalty term. *Inverse Probl.*, 24(5):055020, 13, 2008.
- [10] M. Grasmair, M. Haltmeier, and O. Scherzer. The residual method for regularizing ill-posed problems. *Appl. Math. Comput.*, 218(6):2693–2710, 2011.
- [11] C. W. Groetsch. *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*. Pitman, Boston, 1984.
- [12] T. Hein. Convergence rates for multi-parameter regularization in Banach spaces. *Int. J. Pure Appl. Math.*, 43(4):593–614, 2008.
- [13] T. Hein. Tikhonov regularization in Banach spaces: improved convergence rates results. *Inverse Problems*, 25(3):035002 (18pp), 2009.
- [14] T. Hein and B. Hofmann. Approximate source conditions for nonlinear ill-posed problems – chances and limitations. *Inverse Probl.*, 25:035003, 2009.
- [15] B. Hofmann. Approximate source conditions in Tikhonov-Phillips regularization and consequences for inverse problems with multiplication operators. *Math. Methods Appl. Sci.*, 29(3):351–371, 2006.
- [16] B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer. A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Probl.*, 23(3):987–1010, 2007.
- [17] A. Neubauer. Tikhonov regularization for non-linear ill-posed problems: optimal convergence rates and finite-dimensional approximation. *Inverse Probl.*, 5(4):541–557, 1989.
- [18] C. Pöschl. *Tikhonov Regularization with General Residual Term*. PhD thesis, University of Innsbruck, Austria, Innsbruck, October 2008.
- [19] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, New York, 2009.